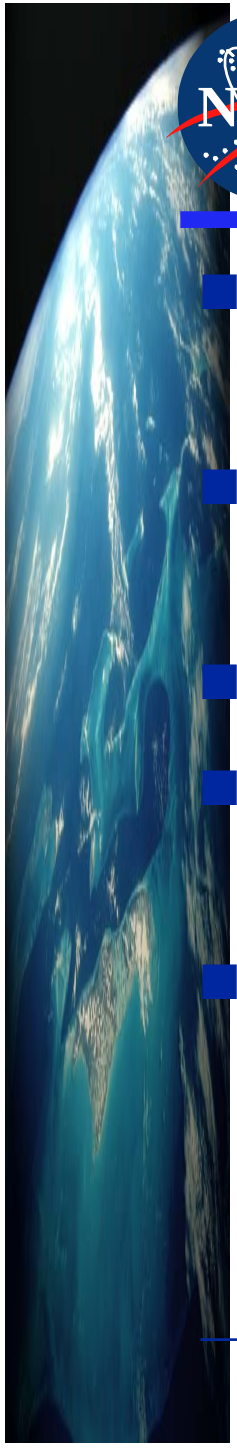
Improving Space Project Cost Estimating With Engineering Management Variables

Joe Hamaker, NASA Headquarters, Washington DC

2nd NASA Project Management Challenge Conference
University of Maryland
March 22-23, 2005



Background and Objectives

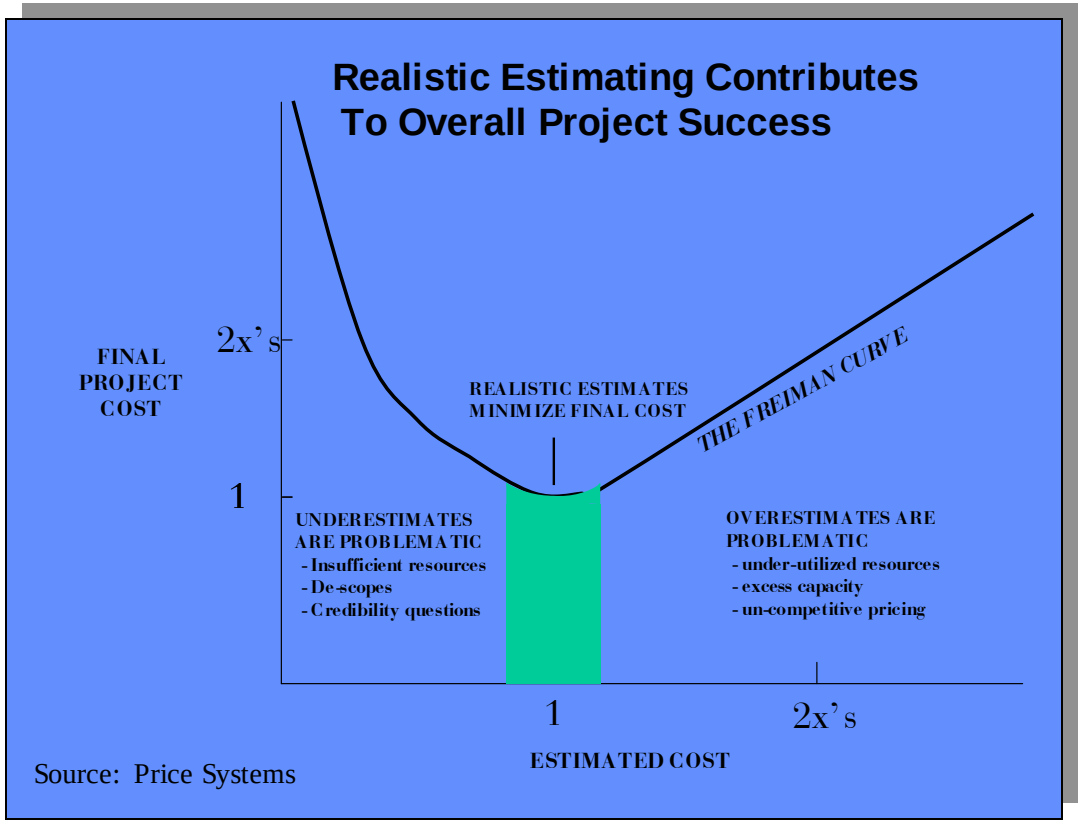
- The aerospace industry uses parametric cost models to predict the cost of future projects
- Current models, typically focus on technical cost variables, using multivariable cost estimating relationships
- Such models leave much variability unexplained
- Common wisdom suggests that the residual cost variability is due to "the way they were managed"
- The objective of this research project is to attempt to improve the models by introducing EM variables

Example Models Using Technical Variables Only

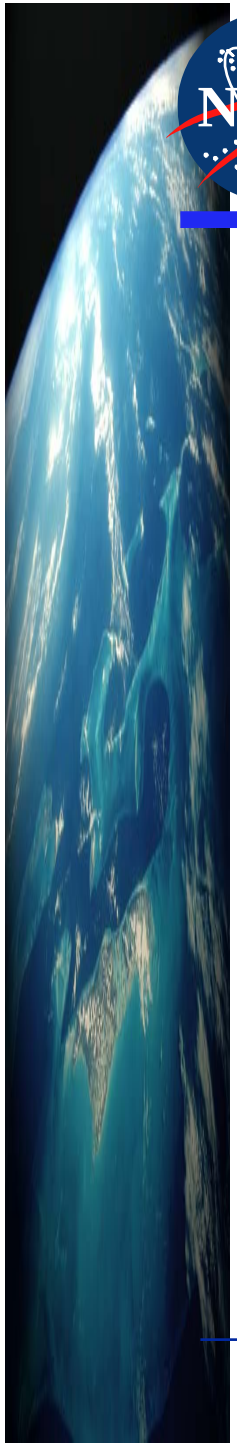
	Significance (p value) In Table			
Regression:	1	2	3	4
Constant	0.001	0.028	0.006	0.064
Dry Mass (kg)	0.000	0.000	0.000	0.000
Max Power (Watts)	--	0.001	0.000	0.000
Max Data Rate (bps)	--	--	0.025	0.055
Total Thrust (Newtons)	--	--	--	0.050
R Squared Adjusted	64.50%	67.10%	68.10%	68.80%

Importance

- The initial budgets of most space projects are set by using current parametric cost models
- If cost models could be improved, more accurate budgeting could be accomplished
- Frank Freiman, the inventor of the Price Model, maintained that realistic estimating contributes to project success
- Additionally, the existence of management variables in cost models provides PMs with feedback on management decisions

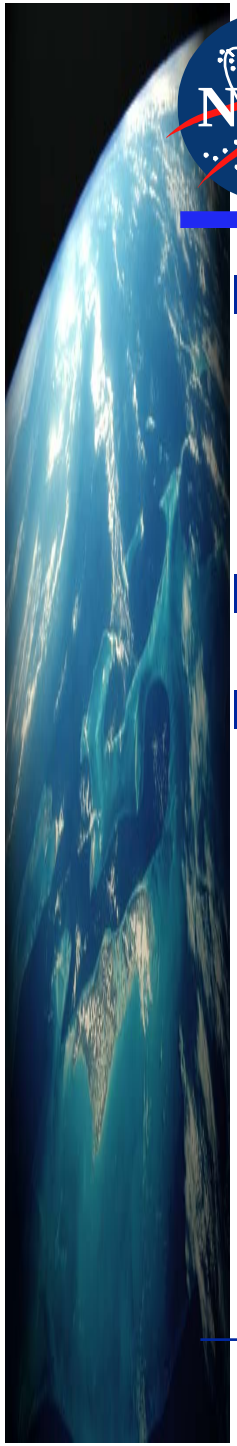


The "Freiman Curve"



Hypothesis

- H_o: Management not shown to have **significant effect** on cost
- H_1: Management shown to have **significant effect** on cost
- At least 3 outcomes possible
  - H_o accepted (Management influence not uncovered by this study)
  - H_o rejected (Management influence seen in this study)
  - Indeterminate--some evidence but statistical analysis could not conclusively show it (signal to noise ratio)



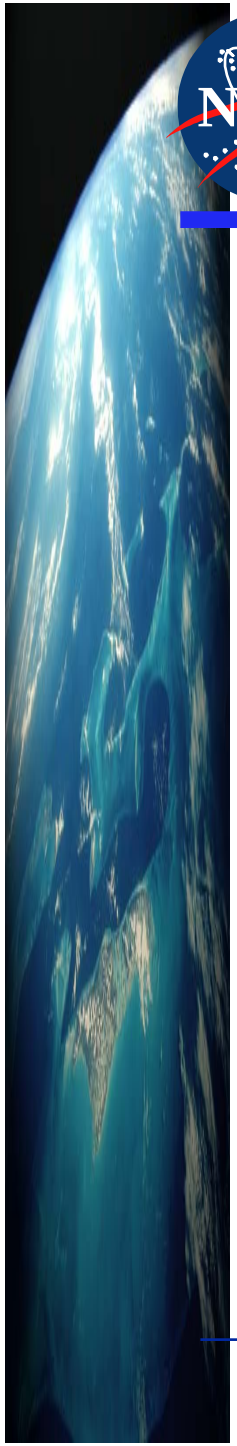
Research To Date

- **To date, 189 projects* have been identified as potential data points**
  - All automated earth orbital and planetary spacecraft
  - Launch dates from July 1964 to December 2009
  - Costs from $3.8 M to $3.8B
- **149 of these projects have been researched**
  - Data tabulated in Excel data base
- **Data sources included**
  - NASA Air Force Cost Model (NAFCOM) data base
  - NASA CFO library and office files
  - NASA Independent Program Assessment Office
  - Project web sites
  - NASA HQ Science Mission Directorate
  - NASA HQ Library
  - Aerospace Corporation
  - Rand Corporation
  - Miscellaneous other sources

Example Projects (Launch Year)

- Cassini (1997)
- Chandra (1999)
- DSCS series (1971-1987)
- Galileo (1989)
- Genesis (2001)
- GP-B (2004)
- Hubble (1990)
- Magellan (1989)
- Mars Pathfinder (1996)
- Mars Exploration Rovers (2003)
- Messenger (2004)
- OSO-8 (1975)
- TRMM (1997)
- Viking (1975)

*Mostly NASA, but some by NOAA, USAF, Navy, DARPA, BMDO, Comsat



Research To Date

- For the 149 projects researched to date, an Excel database has been constructed with 81 data fields
  - 54 of these 81 data fields are thought to be potential independent variables
  - 1 field, total acquisition cost, is the dependent variable
- Independent variables classified into one of three categories (solely for structural convenience)
  - Requirements variables
  - Design variables
  - Management variables
- Approximately 70% of the fields have been populated

Some Examples Of The 81 Database Fields

Requirements Variables

- Development span (months)
- Apogee (km)
- Design Life (months)
- Data rate (kbps)
- Number of communication bands (#)
- Avionics redundancy (categorical)
- Autonomy degree (categorical)
- Pointing accuracy (degrees)
- Power (Watts)
- Amp hours (categorical)
- Number of instruments (#)
- Instrument power (watts)

Management Variables

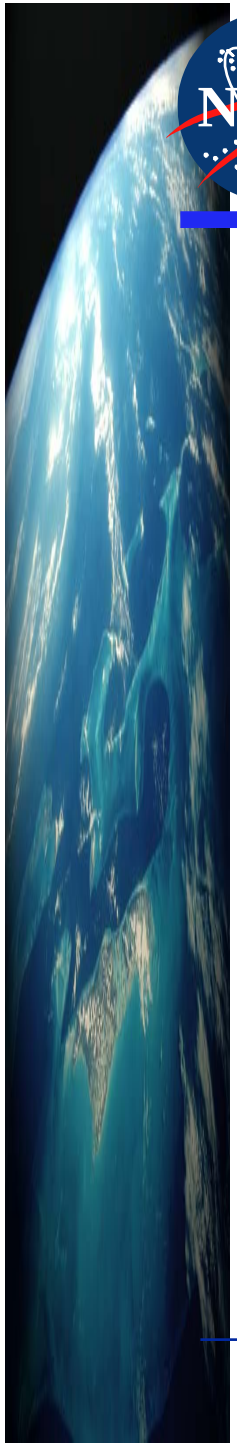
- Acquisition approach (categorical)
- Number of customers (#)
- Number of primes excluding science (#)
- Budget capped (categorical)
- International involvement (categorical)
- Civil service involvement (person years)
- Geographical distribution of partners (categorical)
- Team experience (categorical)
- Team stability (categorical)
- Government lab in-house approach (categorical)
- Degree of prototyping (categorical)
- Requirements volatility (categorical)
- Funding stability (categorical)
- Percent cost spent in formulation (%)
- Percent cost spent on management, systems engineering(%)

Design Variables

- Cylindrical volume (meters^3)
- Total dry mass (kg)
- Instrument dry mass (kg)
- Materials (categorical)
- Number of deployables (#)
- Complexity of deployables (categorical)
- Thermal type (categorical)
- Stability type (categorical)
- GN&C sensors type (categorical)
- Flight software (SLOC)
- Power generation type (categorical)
- Solar array area (meters^2)
- Battery type (categorical)
- Voltage (volts)
- Thrust (Newtons)
- RCS type (categorical)
- Number of RCS thrusters (#)
- Number of science organizations (#)
- Propellant mass (kg)

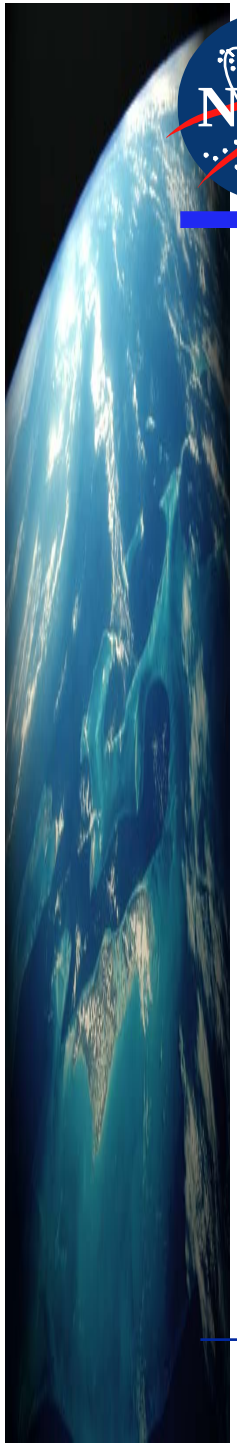
Dependent Variable

- Total Acquisition cost normalized to 100% new design



Some Examples of Categorical Variables

- **Launcher**
  - 0=Expendable launch vehicle
  - 1=Shuttle
- **Materials**
  - 1=Mostly Al
  - 2=Minor composites or exotic metallics
  - 3=Significant composites or exotic metallics
- **Complexity of deployables**
  - 0=None
  - 1=Deploy only
  - 2=Deploy only but complex
  - 3=Deploy and retract
- **Power Generation**
  - 0=Batteries only
  - 1=Silicon solar array
  - 2=GaAs array
  - 3=Fuel cells
  - 4=RTGs
- **Stabilization**
  - 0=None
  - 1=Spin or GG
  - 2=Despun section
  - 3=3 axis controlled with reaction control thrusters
  - 4=3 axis controlled with momentum wheels or reaction wheels
  - 5=3 Axis with Control Moment Gyros (CMGs)
- **Guidance, Navigation & Control Sensors**
  - 0=None
  - 1= Sun sensors
  - 2= earth or horizon sensors
  - 3=star trackers



Allocated Codes Used in Regressions (to Date)

- The categorical variables are included in the analysis using dummy variables (indicator variables)
- As a temporary expediency to obtain a prototype model, the allocated codes on the previous chart have been used in the regressions
- Allocated codes force an undesirable constant spacing of (in this case dollars) for each level of the code
  - For example, using...
  - Guidance, Navigation & Control Sensors
    - 0=None
    - 1=Sun sensors
    - 2=Earth or horizon sensors
    - 3=Star trackers
  - ...forces some constant dollar difference between these four levels of GN&C methods when there is no reason to believe that the levels would have a constant dollar spacing
- Future iterations of the models will introduce n-1 dummy variables to represent n levels, each taking on the value of 0 or 1

Database Status

- Requirements and Design data is relatively readily available
- Some Management data fields are proving to be especially problematic to research
  - Team experience
  - Team stability
  - Requirements stability
- Some Management data fields require more time to research seem "doable" with more persistence
  - Civil service person years
  - Funding stability
  - Percent cost spent on management, systems engineering and integration and test

Data Analysis

- Several prototype models have been developed on the existing data base that show some promise*
- **CostP = 0.0665 + 0.495 DryMassP - 0.00639 TechYr + 0.0385 Apogee + 0.00400 NoDepComp + 0.112 DataP + 0.0905 Stab + 0.137 WattsP + 0.220 CostCap + 0.0501 Acquistion**

Predictor	Coef	StDev	T	P
Constant	0.06646	0.03990	1.67	0.098
DryMass	0.49532	0.06425	7.71	0.000
TechYr	-0.00639	0.001101	-5.80	0.000
Apogee	0.03851	0.01037	3.71	0.000
NoDepComp	0.004004	0.001735	2.31	0.023
DataRateP	0.11165	0.04751	2.35	0.020
StabilityType	0.09054	0.05058	1.79	0.076
WattsP	0.13674	0.05816	2.35	0.020
CostCap	0.21979	0.04744	4.63	0.000
Acquisition	0.05009	0.02377	2.11	0.037

S = 0.1209 R-Sq = 84.0% R-Sq(adj) = 82.8%

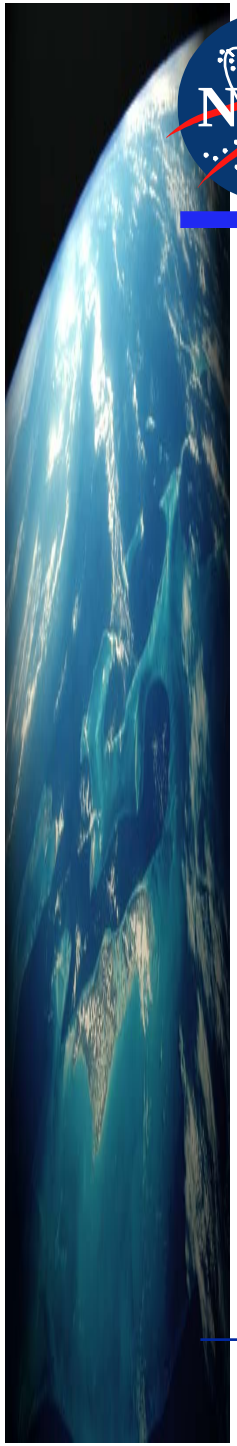
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	9	9.8681	1.0965	75.00	0.000
Residual Error	129	1.8859	0.0146		
Total	138	11.7540			

Decryption of Variables:

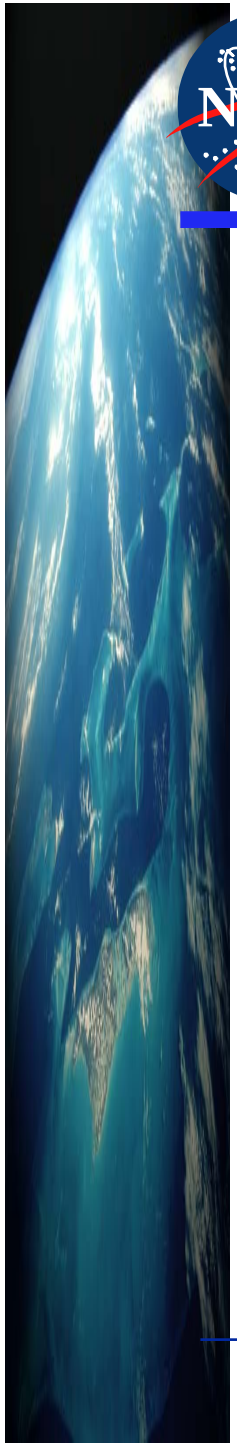
- CostP=Mission cost transformed to a percentile
- DryMassP=Mission total dry mass transformed to a percentile
- TechYr=Year of launch as a proxy for technology date
- Apogee=indicator low earth orbit, high earth orbit or planetary mission
- NoDepComp=Number of deployables multiplied by deployable complexity
- DataRateP= Mission highest data rate transformed to a percentile
- WattsP= Mission power transformed to a percentile
- CostCap=indicator of whether or not mission was cost capped
- Acquisition=indicator for type of acquisition approach used

***The model has an 80% trimmed mean residual error of <12% on the 149 data points**



Areas For Immediate Future Work

- Complete research and population of data base fields
- Investigate influence of outliers (e.g. Hubble, Viking, Surveyor, others)
- Perform more rigorous regression analyses
- Develop workable model(s)
- Validate model both by goodness of fit statistics and by exercising model on projects outside the data base
- Ideally, versions of the model can be developed that work at various points in the project life cycle
  - A version for use early when only Requirements are known
  - A version for use when early Design variables are known
  - But all versions containing Management variables



End